Regular Paper

Proc. of Int. Conf. on Recent Trends in Information, Telecommunication and Computing 2013

Off-line English Character Recognition: A Comparative Survey

Indrani Bhattacherjee[1],

[1]Student, M-Tech (IT), USICT, GGSIPU, Delhi, INDIA

Email:indrani_72@yahoo.com, bhattacherjee.indrani@gmail.com

Abstract: It has been decades since the evolution of idea that human brain can be mimicked by artificial neuron like mathematical structures. Till date, the development of this endeavor has not reached the threshold of excellence. Neural networks are commonly used to solve sample-recognition problems. One of these is character recognition. The solution of this problem is one of the easier implementations of neural networks. This paper presents a detailed comparative literature survey on the research accomplished for the last few decades. The comparative literature review will help us understand the platform on which we stand today to achieve the highest efficiency in terms of Character Recognition accuracy as well as computational resource and cost.

Index Terms- Feature Extraction, Multi-Layered Modular Neural Network, Optical Character Recognition, Pre-Processing.

I. Introduction

Optical Character Recognition, generally abbreviated as OCR, is referred to as the conversion technique of hand written text, typed or digitized text into machine encoded text. Optical Character Recognition (OCR) and Handwritten Character Recognition (HCR) is a part of off-line character recognition. The functionality of OCR lies in input to the system by means of digitized text or hand written text, computational processing of the image to recognize the text successfully.

Although research in the field of Optical Character Recognition has been going on for the last few decades, success in the truest sense has not been totally achievable by the scientists and the goal is still out of reach. Most of the researchers have tried to solve the problem of Optical Character Recognition by means of image processing and pattern recognition techniques. This research has led to the generation of several algorithms for classifications using the rough representation-in-pixels-of the character or feature vector representation.

OCR consists of three foremost features:

- **Pre-processing Stage:** The pre-processing stage is accountable for producing a clean character image to be used directly and efficiently by the feature extraction stage.
- **Feature Extraction Stage:** The feature extraction stage contributes to removing redundancy from data
- **Classification Stage:** The classification stage recognizes characters and words from the algorithm

II. Literature Review

A lot of research has been done in the past few decades on the various methods of character recognition approach with the help of different kinds of artificial neural network, genetic algorithm etc. There are numerous aspects and components of an Optical Character Recognition algorithm that contributes towards a perfect recognition of hand written or typed text input.

Nasien et al. [1] have proposed a recognition model for English handwritten (lowercase, uppercase and letter) character recognition that uses Freeman chain code (FCC) as the representation technique of an image character. Support vector machine (SVM) has been chosen for the classification step. The proposed recognition model, built from SVM classifiers was efficient enough to show that applying the proposed model, a relatively higher accuracy of 98.7% for the problem of English handwritten recognition was reached.

Fuliang et al. [2] proposed that according to the characteristics of vehicle license plate, recognition algorithm could be adapted based on back propagation (BP) neural network. This neural network design could effectively simplify the network structure, improved recognition accuracy and speed. BP algorithm went along improvement as the defects of the standard BP algorithm which had slow convergence and easy to fall into local minimum points. The test results of 100 test samples showed that the whole recognition rate of the character recognition system was 96%, recognition speeding was 301ms.

Deng et al. [3] proposed in their work target detection and pattern recognition as a kind of communications problem and applies error-correcting coding to the outputs of a convolutional neural network to improve the accuracy and reliability of detection and recognition of targets. The outputs of the convolutional neural network were designed according to codewords with maximum Hamming distances. The reliability obtained for isolated hand written digits was around 99.6% - 99.7%.

Gupta et al. [4] focused especially on ofûine recognition of handwritten English words by ûrst detecting individual characters. The main approaches for ofûine handwritten word recognition could be divided into two classes, holistic and segmentation based. Three networks have been considered: Multi-layer perceptron (MLP), radial basis function (RBF) and support vector machine (SVM). The validation yielded poor results for Multi-layer Perceptron Network (MLP). In

DOI: 03.LSCS.2013.4.107

the case of the SVM, the recognition rate on the training data is 98.86% and it achieves the optimum learning. The recognition result on the test data is 62.93%. It is observed that on the test data SVM outperforms the other two networks.

Rashid et al. [5] proposed a segmentation free text line recognition approach using multi-layer perceptron (MLP) and Hidden Markov Models (HMMs). A line scanning neural network trained with character level contextual information and a special garbage class was used to extract class probabilities at every pixel succession. In evaluations on a subset of UNLV-ISRI document collection, 98.4% character recognition accuracy was achieved that was statistically signiûcantly better in comparison with character recognition accuracies obtained from state-of-the-art open source OCR systems.

Wang et al. [6] used generalized regression neural network (GRNN) in character recognition and did some research in license plate recognition system. Generalized regression neural network (GRNN) structure with the advantages of simple design, fast convergence speed required less training samples, the modeling of a prior knowledge of the objects that do not require much, with global approximation and the best approximation property, robustness and strong nonlinear processing ability, according to the sample data reflect the implicit mapping relationship, and no local minimum problem. The method had good performance of ratio in character recognition (around 95.5%). But there was more effort on improving the ratio of recognition so as to apply it into actual license plate recognition system.

Huang et al. [7] presented a neural network based approach to largely reduce the training time while maintain the high recognition rate. The main idea was to perform a preprocessing stage on the training data prior to the neural network training and use a template matching technique in the recognition stage. This algorithm yielded a recognition error rate of 3.05-5% with a high computational cost.

Shrivastava et al. [8] have described in their experiments the performance evaluation for the feed forward neural network with three different soft computing techniques to recognize hand written English alphabets. Numerous potential in the field of pattern recognition have been shown by evolutionary algorithms for the hybrid neural network. It could be clearly understood from their results that there is large significant difference between the performance of back propagation algorithm, evolutionary algorithm (Genetic algorithm) and hybrid evolutionary algorithm. This comparison had been made on the basis of number of iteration, efficiency and rate of convergence. The results indicate that the performance of hybrid evolutionary was better from both the algorithms in terms of convergence and efficiency.

Steinherz et al. [9] presented a novel loop modeling and contour-based handwriting analysis that improves loop investigation. We show excellent results on various loop resolution scenarios, including axial loop understanding and collapsed loop recovery. An approach for loop investigation on several realistic data sets of static binary images was demonstrated and compared with the ground truth of the genuine online signal. In Encapsulated “Hole” Classification experiment, given 259 of 287 authentic natural sub loops (90.2 percent) were successfully detected, false alarms happened in 18 of 253 (7.1 percent) instances, where authentic artificial or superfluous “holes” were mistakenly labeled as natural. This produced a total “hole” identification rate of 91.5 percent (494/540).

Azzopardi et al. [10] proposed a trainable filter called Combination of Shifted Filter Responses (COSFIRE) which was used for key point detection and pattern recognition. It was automatically configured to be selective for a local contour pattern specified by an example. The configuration comprised selecting given channels of a bank of Gabor filters and determining certain blur and shift parameters. . The proposed COSFIRE filters provided effective machine vision solutions in three practical applications: the detection of vascular bifurcations in retinal fundus images (98.50 percent recall and 96.09 percent precision), the recognition of handwritten digits (99.48 percent correct classification), and the detection and recognition of traffic signs in complex scenes (100 percent recall and precision).

Papavassiliou et al. [11] presented two novel approaches to extract text lines and words from handwritten document. The line segmentation algorithm was based on locating the optimal succession of text and gap areas within vertical zones by applying Viterbi algorithm. Then, a text-line separator drawing technique was applied and finally the connected components were assigned to text lines. An accepted threshold was set to 95% and 90% for line and word detection respectively. In line segmentation, the document image was divided in vertical zones and the extreme points of the piece-wise projection profiles were used to over-segment each zone in “gap” and “text” regions.

Pirlo et al. [12] introduced a new class of zone-based membership functions with adaptive capabilities and showed its effectiveness. The basic idea was to select, for each zone of the zoning method, the membership function best suited to exploit the characteristics of the feature distribution of that zone. In addition, a genetic algorithm was proposed to determine—in a unique process—the most favorable membership functions along with the optimal zoning topology. The problem of membership function selection for zoning-based classification in the context of handwritten numeral and character recognition was successfully addressed. A recognition rate of around 99% was shown by this technology.

III. Comparison between Literature Surveys

From the literature survey, it has been studied that researchers have tried out different algorithms for increasing the accuracy of the Optical Character Recognition technique. Out of the all the methods the HMM Models and SVM models have contributed to the highest level of accuracy. With a better strategized method of hybridization technique, one may achieve even a better accuracy range.

DOI: 03.LSCS.2013.4.107

IV. Proposed Algorithm: "Off-Line Handwritten English Character Recognition Using Modular Multi-Layered Neural Networks"

The primary learning from the literature survey was the comparative study of the different types of algorithms, the comparative accuracy of Optical Character Recognition and the computational time.

Many existing character-recognition machines are designed to make a decision on a present character on the basis of measurements on this character alone, without using any information. Handwritten image normalization from a scanned image includes several steps, which usually begin with image cleaning, page skew correction, and line detection. After the slope correction, slant is removed by means of a two-step method. In the first step, a global slant angle is estimated and removed by performing a shear operation to the image for every integer angle between intervals.

When the image is slope and slant-corrected, the size of the text line is normalized in order to minimize the variations in size and position of its three zones (main body area, ascenders, and descenders). Furthermore, the normalized size of ascenders and descenders is reduced with respect to the body since they are not as informative (the presence or absence of ascenders and descenders is preserved, as well as the width, but not the actual height).

After preprocessing, a feature extraction method is applied to capture the most relevant characteristics of the character to recognize. In our system, a handwritten text line image is converted into a sequence of fixed-dimension feature vectors. Following [10], features are extracted by applying a grid to the image and computing three values for each cell of the grid: the normalized gray level and the horizontal and vertical gray level derivatives. A grid of square cells with 20 rows has been used.

In the proposed network architecture, the preprocessed characters are arranged in 16 x 16 bitmap format and serve as input to the multilayered modular network.

The input bitmap is connected locally to a hidden layer of 2704 (52 x 52) hidden nodes. The connection scheme between the input and the first hidden layer of this net is local with a window size of 4 x 4 and with a moving increment of 2 pixels. For recognizing characters, there are 52 small independent subnets, each of which is responsible for a particular character. Each of the subnet has 2 hidden layers and 1 output layer. Here, decisions are made about the correct output for the entire network on a winner takes all method.

The output of the locally connected layer is connected fully to the first hidden layer of the subnet which consists of 208 (52 x 4) nodes.

Input values are summed as followed:

$$A_i = \sum_{j=1}^{2704} w_{ij} o_j \tag{1}$$

where, w_{ij} is the weight values from the ith node in the upper layer to the jth in the lower layer and o_j is the output of node j of the locally connected layer. These values are mapped to activation values of the hidden layer using the standard sigmoid function:

$$o_i = \frac{1}{1+e^{-A_i}} \tag{2}$$

Each node in the first hidden layer of the subnet is fully connected to the second hidden layer; each of these layers consists of 104 (52 x 2) nodes. The full connection approach was preferred over local or shared weight connection scheme for the last hidden layer because experiments with the latter approach did not affect the overall system accuracy by more than half a percent

The second hidden layer of the subnet is fully connected to the third (i.e. the output) layer which consists only of two nodes. The first node plays an important role and its activation represents the recognition of the corresponding class of the subnet. The other node is the complement node, whose activation represents the recognition of a class that does not belong to the subnet. The 52 different subnets yield a set of 104output nodes which provide the output vector used for classification of the input bitmap.

A supervised training algorithm has been used for training the network.

V. Simulation and Results Of Proposed Algorithm

A. Simulation

The simulation has been done on the basis of recognizing characters both uppercase and lowercase. The whole program operates through a MATLAB GUI (Graphical User Interface) which provides the facility of image processing and training through neural networks.

Off-Line Handwritten English character recognition is based on 3 main steps: 1.Image processing 2.Training the characters with modular multilayered neural network 3.Retrieving the characters as a correspondence of training and image processing..

A.1.ImageProcessing

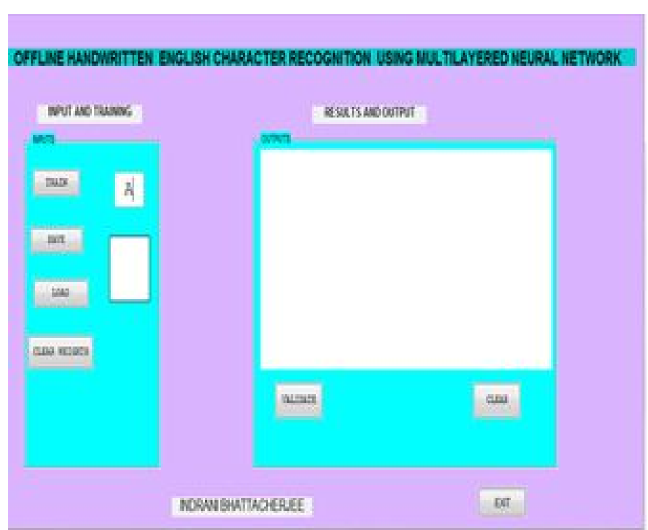

Figure1: GUI for Offline Optical Recognition

DOI: 03.LSCS.2013.4.107

Regular Paper

Proc. of Int. Conf. on Recent Trends in Information, Telecommunication and Computing 2013

Figure2: Input by Hand Written Characters

As seen in the GUI, the bitmap image retrieval box as marked in red shows the binary image through image processing.

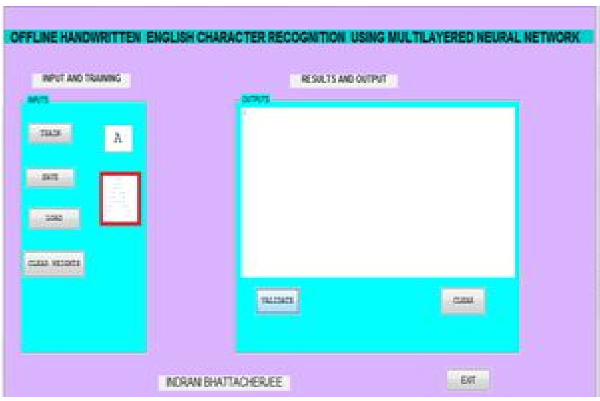

Figure3: Red Marked Bitmap Image Retrieval Box

The bitmap image that comes by image processing of the character A is given as follows:

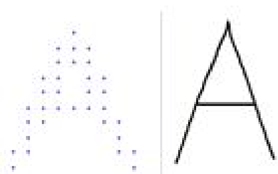

Fig 4: Bitmap Image of character A

A.2 Training

The training has been done on the basis of modular multi-layered neural network

A.3 Retrieval Section

Retrieval section displays the recognized character.

Figure5: Validation Section

B. Results

The no of training samples were 136592.

The amount of errors and the accuracy rates are as follows:

Table: I. Accuracy Table for Capital characters and Small characters

Character	No of Samples	Errors
A	2228	4
B	310	3
C	982	8
D	583	8
E	1345	7
F	5674	4
G	4533	5
H	6574	8
I	564	5
J	6543	11
K	324	4
L	2354	2
M	6334	5
N	675	2
O	432	5
P	435	5
Q	3546	8
R	3425	3
S	987	2
T	213	2
U	785	2
V	6526	14
W	356	3
X	4567	5
Y	6546	13
Z	4367	2
Total	71208	140
Characters	**Samples**	**Errors**
a	3456	5
b	345	4
c	5453	7
d	567	5
e	324	3
f	4566	4
g	435	3
h	4545	5
i	387	3
j	988	3
k	976	9
l	7567	12
m	546	5
n	4576	7
o	545	6
p	6777	8
q	545	4
r	5477	6

DOI: 03.LSCS.2013.4.107

s	1324	5
t	9876	7
u	435	4
v	765	2
w	456	4
x	3542	5
y	344	2
z	567	2
Total	65384	130

The errors for capital letters are 0.196% and for small letters is 0.198%. The total error of OCR is 0.197%. The total accuracy of OCR by modular multi-layered is 99.80%.

The factor of noise is the main factor for the amount of errors in the algorithm. This can avoided by incorporated the algorithm for intelligent removal of errors.

V. Conclusion and Scope For Future Work

The above simulations conclude that the algorithm needs to be more perfect in terms of incorporating the noise factor. Then only a perfect algorithm can be brought down for a perfect implementation for vehicle number selection, cheque identification or advanced string recognition as a part of OCR.

References

[1] Dewi Nasien, Habibollah Haron & Siti Sophiayati Yuhaniz, Support Vector Machine (SVM) for English Handwritten Character Recognition, Second International Conference on Computer Engineering and Applications, 2010

[2] Li Fuliang, Gao Shuangxi, Character Recognition System Based on Back-propagation Neural Network, International Conference on Machine Vision and Human-machine Interface,2010.

[3] Huiqun Deng, George Stathopoulos, & Ching Y. Suen, Applying Error-Correcting Output Coding to Enhance Convolutional Neural Network for Target Detection and Pattern Recognition, International Conference on Pattern Recognition, 2010.

[4] Anshul Gupta, Manisha Srivastava & Chitralekha Mahanta, Ofûine Handwritten Character Recognition Using Neural Network, International Conference on Computer Applications and Industrial Electronics, 2011

[5] Sheikh Faisal Rashid, Faisal Shafait & Thomas M. Breuel, 10th IAPR International Workshop on Document Analysis Systems, 2012

[6] Shenghui Wang & Fei Yang, Research in Character Recognition based on Generalized Regression Neural Network, International Conference on Computing, Measurement, Control and Sensor Network, 2012.

[7] Nan-Chi Huang and Huei-Yung Lin, A Multi-Stage Processing Technique for Character Recognition, IEEE/ASME International Conference on Advanced Intelligent Mechatronics, 2012.

[8] Saurabh Shrivastava, Manu Pratap Singh, Performance evaluation of feed-forward neural network with soft computing techniques for hand written English alphabets, Applied Soft Computing, 2010, 1156-1182.

[9] Tal Steinherz, David Doermann, Ehud Rivlin & Nathan Intrator, Offline Loop Investigation for Handwriting Analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 2, February 2009

[10] George Azzopardi and Nicolai Petkov, Trainable COSFIRE Filters for Keypoint Detection and Pattern Recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 35, No. 2, February 2013

[11] Vassilis Papavassiliou, Themos Stafylakis, Vassilis Katsouros & George Carayannis, Handwritten document image segmentation into text lines and words, Pattern Recognition, 43, 369 – 377, 2010

[12] Giuseppe Pirlo and Donato Impedovo, Adaptive Membership Functions for Handwritten Character Recognition by Voronoi-Based Image Zoning, IEEE Transactions on Image Processing, Vol. 21, No. 9, September 2012

DOI: 03.LSCS.2013.4.107